Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/DK04/000896

International filing date: 22 December 2004 (22.12.2004)

Document type: Certified copy of priority document

Document details: Country/Office: DK
Number: PA 2004 00021
Filing date: 08 January 2004 (08.01.2004)

Date of receipt at the International Bureau: 14 January 2005 (14.01.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

PCT/DK 2004/000896

Kongeriget Danmark

Patent application No.:	PA 2004 00021
Date of filing:	08 January 2004
Applicant: (Name and address)	Novozymes A/S Krogshøjvej 36 DK-2880 Bagsværd Denmark

Title: Amylase

IPC: -

This is to certify that the attached documents are exact copies of the above mentioned patent application as originally filed.

Patent- og Varemærkestyrelsen
Økonomi- og Erhvervsministeriet

07 January 2005

Susanne Morsing

AMYLASE

FIELD OF THE INVENTION

The present invention relates to a polypeptide having amylase activity, to DNA encoding it and to its use in the preparation of dough and dough-based products.

BACKGROUND OF THE INVENTION

Pdb file 6taa (available at www.rcsb.org) describes a fungal amylase (Taka amylase) from *Aspergillus oryzae*.

SUMMARY OF THE INVENTION

The inventors have identified an amylase from a fungal strain of *Chaetomium sp.* and found that the amylase can increase the shelf life of baked products. More specifically, the amylase in combination with a maltogenic amylase further improves the softness of bread crumb without having detrimental effects on elasticity.

Accordingly, the invention provides a polypeptide having amylase activity. It may be a polypeptide encoded by the amylase-encoding part of the DNA sequence inserted into a plasmid present in *E. coli* DSM 16113 or having an amino acid sequence as shown in positions 1-566 of SEQ ID NO 2, or it may be at least 70 % identical to one of these. The polypeptide may also be encoded by a nucleic acid sequence which hybridizes at 55°C with the complementary strand of nucleotides 146-1843 of SEQ ID NO: 1

The invention also provides a polypeptide having an amino acid sequence which can be obtained from the mature polypeptide of SEQ ID NO: 2 by substitution, deletion, and/or insertion of one or more amino acids and a polynucleotide having a sequence that can be derived from SEQ ID NO: 1 by substitution, deletion, and/or insertion of one or more nucleotides.

The invention also provides a polynucleotide encoding the amylase, an expression vector comprising the polynucleotide, a transformed host cell comprising the vector, as well as a method of producing the amylase by cultivating the transformant. The invention further provides a dough composition comprising the amylase, a method of preparing a dough-based product by leavening and heating the dough, e.g. by baking.

DETAILED DESCRIPTION OF THE INVENTION

Genomic DNA source

The source organism of the amylase of the invention is a fungal strain isolated from soil samples collected from Hainan Province, China, in 2002. The strain was classified as be-

longing to Fungi, *Ascomycota*, *Sordariomycetidae*, *Sordariales*, *Chaetomiacae*, *Chaetomium sp.*

The inventors have cloned the gene encoding the polypeptide of the invention from the source organism into a strain of *E. coli* and deposited it under the terms of the Budapest Treaty on 16 December 2003 as DSM 16113 with the DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, D-38124 Braunschweig DE.

Recombinant expression vector

The expression vector of the invention typically includes control sequences encoding a promoter, operator, ribosome binding site, translation initiation signal, and, optionally, a selectable marker, a transcription terminator, a repressor gene or various activator genes. The vector may be an autonomously replicating vector, or it may be integrated into the host cell genome.

Production by cultivation of transformant

The polypeptide of the invention may be produced by transforming a suitable host cell with a DNA sequence encoding the lipolytic enzyme, cultivating the transformed organism under conditions permitting the production of the enzyme, and recovering the enzyme from the culture.

The host organism may particularly be a eukaryotic cell, in particular a fungal cell, such as a yeast cell or a filamentous fungal cell, e.g. a strain of *Aspergillus*, *Fusarium*, *Trichoderma* or *Saccharomyces*, particularly *A. niger*, *A. oryzae*, *F. graminearum* or *S. cerevisiae*.

Hybridization

Suitable conditions for determining hybridization between a nucleotide probe and a homologous DNA or RNA sequence involve presoaking of the filter containing the DNA fragments or RNA to hybridize in 5 x SSC (standard saline citrate) for 10 min, and prehybridization of the filter in a solution of 5 x SSC (Sambrook et al. 1989), 5 x Denhardt's solution (Sambrook et al. 1989), 0 5 % SDS and 100 µg/ml of denatured sonicated salmon sperm DNA (Sambrook et al 1989), followed by hybridization in the same solution containing a random-primed (Feinberg, A. P. and Vogelstein, B. (1983) *Anal. Biochem.* 132:6-13), ^{32}P-dCTP-labeled (specific activity > 1 x 10^9 cpm/µg) probe for 12 hours at approx 45°C. The filter is then washed two times for 30 minutes in 2 x SSC, 0.5 % SDS at a temperature of at least 55°C, more particularly at least 60°C, more particularly at least 65°C, even more particularly at least 70°C, especially at least 75°C. Molecules to which the oligonucleotide probe hybridizes under these conditions may be detected using a x-ray film.

Alignment and identity

The polypeptide and polynucleotide of the invention may have identities to the disclosed sequences of at least 80 %, particularly at least 85 % or at least 90 %, e.g. at least 95 %.

For purposes of the present invention, alignments of sequences and calculation of identity scores may be done using a Needleman-Wunsch alignment (i.e. global alignment), useful for both protein and DNA alignments. The default scoring matrices BLOSUM50 and the identity matrix are used for protein and DNA alignments respectively. The penalty for the first residue in a gap is -12 for proteins and -16 for DNA, while the penalty for additional residues in a gap is -2 for proteins and -4 for DNA Alignment is from the FASTA package version v20u6 (W. R. Pearson and D. J Lipman (1988), "Improved Tools for Biological Sequence Analysis", PNAS 85.2444-2448, and W. R Pearson (1990) "Rapid and Sensitive Sequence Comparison with FASTP and FASTA", Methods in Enzymology, 183:63-98).

Dough

The dough of the invention generally comprises flour, particularly wheat flour. The dough may be fresh, frozen or par-baked. It may be a laminated dough.

The dough may also comprise other conventional dough ingredients, e.g.. proteins, such as milk powder and gluten; eggs (either whole eggs, egg yolks or egg whites); an oxidant such as ascorbic acid, potassium bromate, potassium iodate, azodicarbonamide (ADA) or ammonium persulfate; an amino acid such as L-cysteine; a sugar; a salt such as sodium chloride, calcium acetate, sodium sulfate or calcium sulfate The dough may comprise fat (triglyceride) such as granulated fat or shortening.

Additional enzyme

Optionally, one or more additional enzymes may be added to the dough together with the amylase of the invention. The additional enzyme may be a second amylase, a lipolytic enzyme (e.g. as described in WO 9953769) or a xylanase. The second amylase may be an exo-acting maltogenic alpha-amylase, e.g. as described in WO 9104669 or WO 9943794; an example is Novamyl® (product of Novozymes A/S)

Dough-based product

The invention provides a method for preparing a dough-based product by leavening the dough and heating it, e.g. by baking or steaming. The dough may be leavened e.g. by adding chemical leavening agents or yeast, usually *Saccharomyces cerevisiae* (baker's yeast). The product may be of a soft or a crisp character, either of a white, light or dark type. Exam-

ples are steamed or baked bread (in particular white, whole-meal or rye bread), typically in the form of loaves or rolls.

Amylase Units (AmU)

Amylase activity was assayed by incubating with Phadebas amylase test tablets (product of Amersham Pharmacia) suspended in 50 mM sodium acetate + 1 mM $CaCl_2$ at pH 5.7 and determining OD at 650 nm. The Amylase Unit (AmU) was defined by taking the activity of the commercial product BAN 480L (product of Novozymes A/S) as 480 AmU.

EXAMPLES

Materials and methods

RNeasy Mini Kit (Qiagen, Cat. #74904).

Taq DNA polymerase (Promega, Cat. # M166A)

pGEM-T Vector System I (Promega, Cat. # A3600)

Wizard Plus Minipreps DNA Purification System (Promega, Cat. # A7510)

5' Rapid Amplifiction of cDNA End System (Life Technologies, 5'RACE, Cat # 18374-041),

3' Rapid Amplifiction of cDNA End System (Life Technologies , 3' RACE, Cat. # 1085805)

ElectroMAX DH10B Cells (Life Technologies, Cat. # 18290-015)

Example 1: Cultivation of fungal strain for cDNA preparation

The fungal strain *Chaetomium sp.* was grown on YG agar plate (4.5 cm diam) for 5 days under 37°C in the darkness and used for inoculating shake flask. The plates with fully grown cultures were stored at 4°C before use.

To obtain the mycelium for cDNA library construction, 4-6 agar plugs with fully grown fungal cultures on the YG agar plates were used to inoculate one shake flask with FG-4 (50 ml in 500 ml Erlenmeyer flask with 2 baffles: 30 g Soymeal, 15 g Maltose, 5 g Peptone, 1000 ml H_2O, 1% starch, 1 g olive oil (2 drops / flask), Autoclave at 121 °C for 30 min) and grown under 37°C, 160 rpm for 24 hours. The mycelium was harvested by centrifugation of the culture broth at 8000 rpm and 4°C for 30 minutes. Then mycelium was transferred into a clean plastic bag following by immediately freezing in liquid nitrogen and stored at -80°C before total RNA was isolated.

Probe designing:

Degenerate primers were designed based on alignment of already known amylase gene sequences. amyD1 (SEQ ID NO: 3) and amyD2R (SEQ ID NO: 4).

The N-terminal amino acid sequence of the purified amylase AM835F was determined as shown in SEQ ID NO. 13. This was used to design four degenerate primers: AM835n-s1 (SEQ ID NO: 5), AM835n-s2a (SEQ ID NO: 6), AM835n-s2b (SEQ ID NO: 7), AM835n-s2c (SEQ ID NO: 8).

Extraction of total RNA:

Total RNA was isolated from the frozen mycelium of a strain of *Chaetomium sp.* by using RNeasy Mini Kit according to the manufacturer's instructions.

Gene cloning:

cDNA was synthesized using 3' RACE kit. The primary PCR was performed by using N-terminal based degenerate primers (primer AM835n-s2 is a mixture of AM835n-s2a, b and c) with AUAP provided by the 3'RACE kit:

10X PCR buffer	5 micro-l
25mM MgCl2	3 micro-l
10mM dNTP	1 micro-l
100 micro-M AM835n-s2a	1 micro-l
100 micro-M AM835n-s2b	1 micro-l
100 micro-M AM835n-s2c	1 micro-l
AUAP	1 micro-l
cDNA	2 micro-l
Taq DNA polymerase (5u/ micro-l)	1 micro-l
H_2O	34 micro-l

The PCR program was: 94°C for 3 min; 30 cycles of 94 °C for 30s, 50°C for 30s, 72°C for 1.5 min; final extension at 72°C for 10 min.

There was no specific amplification seen when the PCR product was visualized under UV but this product was used for second PCR with degenerate primers designed based on amylase homology. The 2nd PCR was performed by using amylase probes (amyD1 and amyD2R) and using the primary PCR as template:

10X PCR buffer	5 micro-l
25mM MgCl2	3 micro-l
10mM dNTP	1 micro-l
100 micro-M amyD1	1 micro-l
100 micro-M amy D2R	1 micro-l
1st PCR	1 micro-l

Taq DNA polymerase (5u/ micro-l)	1 micro-l
H_2O	37 micro-l

The PCR program was: 94°C for 3min; 30 cycles of 94 °C for 30s, 50°C for 30s, 72°C for 1min; final extension at 72°C for 10min. A specific band was amplified at ~700bps and this was confirmed to be amylase by sequencing.

Based on the above obtained partial sequence, new primers were designed for 5' and 3' end cloning. For 5' end cloning, after cDNA was synthesized initiated with amy835as1 (SEQ ID NO: 9) by 5'RACE kit, PCR was performed with primer pairs amy835as1 and AAP (provided by the kit). Then nested PCR was performed with primer pair amy835as3 (SEQ ID NO: 14) and AUAP by using primary PCR (amy835as1-AAP) as template. A fragment of ~600bp was obtained and confirmed by sequencing. For 3' end cloning, PCR was performed by using primer pair of amy835f1 (SEQ ID NO: 15) and AUAP and cDNA as template. The nested PCR was performed by using primer pair amy835f2 (SEQ ID NO: 10) with AUAP and 1st PCR as template. A fragment at ~600 bps was amplified and again confirmed by sequencing.

Then based on the cloned 5' and 3' end sequences, the 5' and 3' end primers for full length cloning was designed and used for full length cloning of the amylase AM835. By using cDNA synthesized by 3' RACE kit as template and probes amy835s00 (SEQ ID NO: 11) and amy835as01 (SEQ ID NO: 12) as primers:

10X PCR buffer	5 micro-l
25mM MgCl2	3 micro-l
10mM dNTP	1 micro-l
10 micro-M amy835s00	1 micro-l
10 micro-M amy835as01	1 micro-l
cDNA	2 micro-l
Taq DNA polymerase (5 u/micro-l)	1 micro-l
H_2O	36 micro-l

PCR program was: 94°C for 3 min; 30 cycles of 94 °C for 30s, 50°C for 30s, 72°C for 1.5 min; final extension at 72°C for 10 min.

A specific fragment of ~2.0 kb was PCR-amplified. The fragment was cloned into pGEM-T vector (Promega) which has a 3'-T overhang and transformed into *E.coli* DH10B (ElectroMAX DH10B Cells, available from Life Technologies, Cat. # 18290-015) and further sequenced.

Example 2: Production of amylase

YG and FG-4 media were prepared as follows.

YG: Yeast-glucose agar

5.0 g Difco powdered yeast extract;
20.0 g agar;
10.0 g glucose
1000 ml tap water
Autoclave at 121 °C for 15-20 min.

FG-4 Media 50 ml / flask:

30 g Soymeal,
5 g Peptone,
1 g olive oil (2 drops / flask)
15 g Maltose
1000 ml H_2O
50 ml in 500 ml Erlenmeyer flask with 2 baffles. Autoclave at 121 °C for 30 min

A strain of the thermophilic fungus *Chaetomium sp.* was grown on YG agar plate (4.5 cm diam) for 3 days under 37°C in the darkness and used for inoculating shake flask. The plates with fully grown cultures were stored at 4°C before use.

For enzyme production, 4-6 agar plugs with fully grown fungal cultures on the above plates were used to inoculate one shake flask with FG-4 and grown under 37°C, 160 rpm for 72 hours, then harvested by centrifuged the culture broth at 8000 rpm and 4°C for 30 minutes. The supernatant was collected and used for enzyme purification.

1000 ml supernatant was precipitated with ammonium sulfate (80% saturation) and redissolved in 100 ml 25mM Tris-HCl buffer, pH7.0, then dialyzed against the same buffer and filtered through a 0.45 mm filter, the final volume was 200 ml. The solution was applied to a 35 ml Source 15Q column (Pharmacia) equilibrated in 25 mM Tris-HCl buffer, pH7.0, and the proteins was eluted with a linear NaCl gradient (0 –0.3M). Fractions from the column were analyzed for amylase activity on AZCL-amylose at pH 5.5. Fractions with amylase activity were pooled. Then the pooled solution was ultrafiltrated, the concentrated solution was applied to a 180ml Superdex75 column equilibrated with 25 mM Tris-HCl, pH7.0, the proteins was eluted with the same buffer. Amylase containing fractions were analyzed by SDS-PAGE and pure fractions were pooled.

The purified amylase was used for characterization in the following example.

Example 3: Expression of an amylase from *Chaetomium sp.* in *Aspergillus oryzae*

The DNA sequence of the Chaetomium amylase (SEQ ID NO.: 1) was used to design primers for PCR amplification of the amylase encoding-gene from the clone described in Example 1, with appropriate restriction sites added to the primer ends to facilitate sub-cloning of the PCR product (primers AM835.1 and AM835.2, SEQ ID NO: 16 and 17). PCR amplification was performed using AmpliTaq Gold DNA Polymerase (Applied Biosystems, Foster City, CA, USA) following the manufacturer's instructions and using an annealing temperature of 55°C for the first 5 cycles and 65°C for an additional 25 cycles and an extension time of 2 minutes

The PCR fragment was restricted with *BamHI* and *XhoI* and cloned into the *Aspergillus* expression vector pMStr57 using standard techniques. The expression vector pMStr57 contains the same elements as pCaHj483 (WO 98/00529), with minor modifications made to the *Aspergillus* NA2 promoter as described for the vector pMT2188 in WO 01/12794, and has sequences for selection and propogation in *E. coli*, and selection and expression in *Aspergillus*. Specifically, selection in *Aspergillus* is facilitated by the *amd*S gene of *Aspergillus nidulans*, which allows the use of acetamide as a sole nitrogen source. Expression in *Aspergillus* is mediated by a modified neutral amylase II (NA2) promoter from *Aspergillus niger* which is fused to the 5' leader sequence of the triose phosphate isomerase (tpi) encoding-gene from *Aspergillus nidulans*, and the terminator from the amyloglucosidase-encoding gene from *Aspergillus niger*. The amylase-encoding gene of the resulting *Aspergillus* expression construct, pMStr91, was sequenced and the sequence agreed completely with that determined previously.

The Aspergillus oryzae strain BECh2 (WO 00/39322) was transformed with pMStr91 using standard techniques (Christensen, T et al., (1988), Biotechnology 6, 1419-1422). Transformants were cultured in YP+2%G medium shaken at 250 RPM at 30°C and expression of amylase was monitored by SDS-PAGE.

<u>Medium YP+2%G</u>

10g yeast extract
20g peptone
water to 1L
autoclave at 121°C, 20 minutes
add 100ml 20% sterile glucose solution

Example 4: Characterization of amylase

The molecular weight of the amylase prepared in a previous example was found to be around 66 kDa as seen on SDS-PAGE. The isoelectric point (pI) was found to be around pH 3.5, as determined by isoelectric focusing (IEF).

pH and temperature profiles were determined with AZCL-amylose (product of Megazyme) as substrate. At 50°C, the amylase was found to be active at pH 4-10 with an optimum around pH 5-7. At pH 5.5, the amylase was found to be active at 20-70°C with an optimum around 60°C. Thus, the *Chaetomium* amylase has a wider pH range and a higher temperature optimum than the fungal amylase from *Aspergillus oryzae*.

Stability of the amylase was determined by incubation at pH 5-7 and 60-80°C for 5- 25 minutes. The results showed more than 90 % residual activity after 20 minutes at pH 6-7 and 60°C. At pH 5.0 and 60°C, the amylase was nearly completely inactivated in 15 minutes. At

70°C, the amylase was nearly completely inactivated at in 5-10 minutes at pH 6-7. It was found that at all conditions the *Chaetomium* amylase is more stable than the fungal amylase from *Aspergillus oryzae*

The amylase showed no activity on the following substrates at pH 7.0: AZCL-galactomannan, AZCL-beta-glucan, AZCL-dextran, AZCL-xyloglucan, AZCL-potato galactan, AZCL-arabinan, AZCL-pullulan, AZCL-xylan, AZCL-he-cellulose and AZCL-casein

Example 5: Effect of amylase on freshness of bread

Bread were baked according to the sponge & dough method.

Recipes

Sponge	% on flour basis
Soya oil	2,5
Sodium stearoyl lactylate (SSL)	0,38
Yeast	5
Wheat flour	60
Water	62

Dough	% on flour basis
Ascorbic acid	optimized for each flour
ADA	20 ppm
Salt	2
Syrup	7 (dry substance)
Water	optimized for each flour
Wheat flour	40
Calcium propionate	0 25
Enzymes	as indicated below

Sponge

Scaling of ingredients, addition of yeast, water, flour, SSL and oil into mixer bowl

Mixing 90 rpm for 1 minutes, 150 rpm for 4 minutes

The sponge was weighted, the temperature was measured and the sponge was placed in a bowl ~ fermentation 3 hours at 27 C, 86 % RH

Dough

Addition of ingredients and the sponge into the mixer bowl. The sponge and ingredients were mixed together 90 rpm for 9 minutes

The temperature was measured, dough characteristics were evaluated, the dough was scaled into smaller pieces of 435 g each.

The dough rests on the table for 10 minutes

Doughs were sheeted and molded.

Fermentation for 55 minutes at 42°C and 86% RH.

Bread were baked at 200°C for 22 minutes

Enzymes were dosed at 400 MANU/kg of Novamyl together with 0, 5 or 20 AmU/kg of the amylase of SEQ ID NO: 2 (prepared as in Example 1).

Bread were stored at room temperature until analysis.

Texture and water migration by NMR were measured on day 7, 14 and 21. A small sensory evaluation of softness and moistness was performed on day 21

Results

Firmness of the loaves was measured as described in WO 9953769 The results were as follows:

Novamyl dosage MANU/kg	Amylase of invention AmU/kg	Firmness after 7 days g	Firmness after 14 days g	Firmness after 21 days g
400	0	593	869	1103
400	5	505	814	1000
400	20	480	789	939

Elasticity of the loaves was measured as described in US 6162628. The results were as follows:

Novamyl dosage MANU/kg	Amylase of invention AmU/kg	Elasticity after 7 days %	Elasticity after 14 days %	Elasticity after 21 days %
400	0	50.7	46 5	45.2
400	5	50.1	46.7	44 7
400	20	50.7	47.2	46.0

The data show that the amylase of the invention has a significant effect on firmness in combination with Novamyl, furthermore the elasticity seems to be comparable to or even better than that of Novamyl after 21 days of storage.

The mobility of free water was determined as described by P. L. Chen, Z Long, R. Ruan and T P. Labuza, Nuclear Magnetic Resonance Studies of water Mobility in Bread during Storage Lebensmittel Wissenschaft und Technologie 30, 178-183 (1997). The results were as follows:

Novamyl dosage MANU/kg	Amylase of invention AmU/kg	Free water after 7 days Micro-sec	Free water after 14 days Micro-sec	Free water after 21 days Micro-sec
400	0	7498	6921	6198
400	5	7780	6856	6424
400	20	7945	7004	6618

The data show that the amylase of the invention increases the amount of free water The amount of free water has been described in literature to correlate to moistness of bread crumb.

The ranking from the small sensory evaluation of softness and moistness on day 21 showed the following ranking (MANU/kg of Novamyl + AmU/kg of amylase of invention):

Moistest: 400 MANU + 20 AmU

Second: 400 MANU + 5 AmU

Lowest (least moist): 400 MANU

CLAIMS

1. A polypeptide having amylase activity selected from the group consisting of:

a) a polypeptide encoded by the amylase-encoding part of the DNA sequence inserted into a plasmid present in *E. coli* DSM 16113

b) a polypeptide having an amino acid sequence as shown in positions 1-566 of SEQ ID NO 2,

c) a polypeptide which has at least 70 % identity to the polypeptide defined in (a) or (b),

d) a polypeptide which is encoded by a nucleic acid sequence which hybridizes at 55 °C with the complementary strand of nucleotides 146-1843 of SEQ ID NO: 1.

2. A polynucleotide comprising a sequence selected from the group consisting of:

a) the amylase-encoding sequence inserted into a plasmid present in *E. coli* DSM 16113;

b) nucleotides 146-1843 of SEQ ID NO: 1;

c) a polynucleotide encoding amino acids 1-566 of SEQ ID NO 2;

d) a polynucleotide which encodes a polypeptide having amylase activity and has at least 70 % identity to the polynucleotide of a), b) or c),

e) a nucleic acid sequence which hybridizes at 55°C with the complementary strand of the polynucleotide of a), b) or c),

f) the complementary strand of the polynucleotide of a), b), c), d) or e).

3 A vector comprising the polynucleotide of claim 2 operably linked to one or more control sequences that direct the production of the polypeptide in a suitable host.

4. A transformed host cell comprising the vector of claim 3.

5. A method for producing an amylase, which comprises

a) cultivating the host cell of claim 4 under conditions appropriate for expression of *amylase, and*

b) recovering the amylase.

6. A dough composition which comprises flour and the polypeptide of claim 1.

7 A process for preparing a dough-based product, comprising adding the polypeptide of claim 1 to a dough, leavening, and heating the dough.

8. The process of claim 7 which further comprises adding an exo-acting maltogenic alpha-amylase to the dough

Modtaget

-8 JAN. 2004

PVS

SEQUENCE LISTING

<110> Novozymes A/S

<120> Amylase

<130> 10583.000

<160> 17

<170> PatentIn version 3.2

<210> 1
<211> 1946
<212> DNA
<213> Chaetomium sp.

<220>
<221> CDS
<222> (86)..(1843)

<220>
<221> sig_peptide
<222> (86)..(145)

<220>
<221> mat_peptide
<222> (146)..(1843)

<400> 1

```
tcccggtcat cctctcttgg tccctgccat cctcctgccc cctctgatcc accgcctctt      60

cggtggactc caagacgttg tcagg atg cga tcc ttc ctc gcc ctc tca gcc       112
                            Met Arg Ser Phe Leu Ala Leu Ser Ala
                            -20                 -15

ttg ctg ctg ctg tac ccg ctg cag ctg ctc gcc gcc agc aac tcc gac       160
Leu Leu Leu Leu Tyr Pro Leu Gln Leu Leu Ala Ala Ser Asn Ser Asp
    -10                 -5                  -1  1               5

tgg agg tcc cgc aat atc tac ttt gcc ttg acc gac cgc gtc gcc aat       208
Trp Arg Ser Arg Asn Ile Tyr Phe Ala Leu Thr Asp Arg Val Ala Asn
                10                  15                  20

ccg tcc acc acg acc gca tgt agt gac ctg agc aac tac tgc ggc ggc       256
Pro Ser Thr Thr Thr Ala Cys Ser Asp Leu Ser Asn Tyr Cys Gly Gly
            25                  30                  35

acg tgg agc ggc ctg tcg agc aag ctg gac tac atc caa ggg atg ggc       304
Thr Trp Ser Gly Leu Ser Ser Lys Leu Asp Tyr Ile Gln Gly Met Gly
        40                  45                  50

ttc gat tcc atc tgg att acc ccc gtg gtc gag aac tgc gac ggt ggc       352
Phe Asp Ser Ile Trp Ile Thr Pro Val Val Glu Asn Cys Asp Gly Gly
    55                  60                  65

tac cac ggc tac tgg gcc aag gcg ctc tac aac gtc aac acg aac tac       400
Tyr His Gly Tyr Trp Ala Lys Ala Leu Tyr Asn Val Asn Thr Asn Tyr
70                  75                  80                  85

ggc agt gcg gat gat ctg aag aac ttc gtt gcg gcc gcc cat gcg aag       448
Gly Ser Ala Asp Asp Leu Lys Asn Phe Val Ala Ala Ala His Ala Lys
                90                  95                  100

ggc atg tac gtg atg gtg gac gtc gtc gcg aat cac atg ggt tcc tgc       496
Gly Met Tyr Val Met Val Asp Val Val Ala Asn His Met Gly Ser Cys
            105                 110                 115
```

```
ggc atc gcc aac ctc tcc cca cct ccc ctg aac gag cag agc tct tat        544
Gly Ile Ala Asn Leu Ser Pro Pro Pro Leu Asn Glu Gln Ser Ser Tyr
        120                 125                 130

cac acc cag tgc gac att gac tac agc agt cag tcc agc att gag acg        592
His Thr Gln Cys Asp Ile Asp Tyr Ser Ser Gln Ser Ser Ile Glu Thr
    135                 140                 145

tgc tgg ata tcc ggc ctc cct gac ctg gac acc acc gat agc act atc        640
Cys Trp Ile Ser Gly Leu Pro Asp Leu Asp Thr Thr Asp Ser Thr Ile
150                 155                 160                 165

cga tcc ctc ttc cag acc tgg gtc cac ggc ctg gtc agc aac tac agc        688
Arg Ser Leu Phe Gln Thr Trp Val His Gly Leu Val Ser Asn Tyr Ser
                170                 175                 180

ttc gac ggt ctc cgc gtc gac acc gtc aag cac gtg gag aag gat tac        736
Phe Asp Gly Leu Arg Val Asp Thr Val Lys His Val Glu Lys Asp Tyr
            185                 190                 195

tgg ccc ggc ttc gtg tcg gcg gcg ggc acc tac gcc atc ggc gaa gtc        784
Trp Pro Gly Phe Val Ser Ala Ala Gly Thr Tyr Ala Ile Gly Glu Val
        200                 205                 210

ttc tcc ggc gac acc tcc tac gtg gcc ggc tat caa tcg gtg atg ccg        832
Phe Ser Gly Asp Thr Ser Tyr Val Ala Gly Tyr Gln Ser Val Met Pro
    215                 220                 225

ggc ttg ctc aac tat ccc atc tac tat ccg ctc atc cgc gtc ttc gcg        880
Gly Leu Leu Asn Tyr Pro Ile Tyr Tyr Pro Leu Ile Arg Val Phe Ala
230                 235                 240                 245

cag ggt gcg tcc ttc acc gat ctc gtc aac aac cac gat acc gtc ggc        928
Gln Gly Ala Ser Phe Thr Asp Leu Val Asn Asn His Asp Thr Val Gly
                250                 255                 260

tcg acc ttc tcc gac ccg acg ctg ctg ggt aac ttt atc gac aac cac        976
Ser Thr Phe Ser Asp Pro Thr Leu Leu Gly Asn Phe Ile Asp Asn His
            265                 270                 275

gac aac cca cgt ttc ctg agc tac acc agc gac cac gcc ctc ctc aag       1024
Asp Asn Pro Arg Phe Leu Ser Tyr Thr Ser Asp His Ala Leu Leu Lys
        280                 285                 290

aac gct ctg gcc tac gtc atc ctg gcc aga ggc atc ccc atc gtc tac       1072
Asn Ala Leu Ala Tyr Val Ile Leu Ala Arg Gly Ile Pro Ile Val Tyr
    295                 300                 305

tac ggc acc gag caa ggc tac tcg ggt tcg tcc gac ccg gcg aac cgc       1120
Tyr Gly Thr Glu Gln Gly Tyr Ser Gly Ser Ser Asp Pro Ala Asn Arg
310                 315                 320                 325

gag gat ctc tgg cgt agc gga tac agc act acg gga gac atc tac acc       1168
Glu Asp Leu Trp Arg Ser Gly Tyr Ser Thr Thr Gly Asp Ile Tyr Thr
                330                 335                 340

acc atc gcc gcg ctc tcc gcc gcg cgc acc gcg gcc ggt ggc ctc gcc       1216
Thr Ile Ala Ala Leu Ser Ala Ala Arg Thr Ala Ala Gly Gly Leu Ala
            345                 350                 355

ggt aac gac cac gtc cac ctg tac acg acc gac aac gcg tac gcc tgg       1264
Gly Asn Asp His Val His Leu Tyr Thr Thr Asp Asn Ala Tyr Ala Trp
        360                 365                 370

tcc cgg gcg agc ggc aag ctc atc gtc gtc acg tcc aac cgc ggc agc       1312
Ser Arg Ala Ser Gly Lys Leu Ile Val Val Thr Ser Asn Arg Gly Ser
    375                 380                 385
```

```
tcc gac agc agc acc atc tgc ttc agc acc cag cag gcc agc ggc acc      1360
Ser Asp Ser Ser Thr Ile Cys Phe Ser Thr Gln Gln Ala Ser Gly Thr
390                 395                 400                 405

acc tgg acc agc acg atc acc ggc aac tcg tac acc gcc gac agc aac      1408
Thr Trp Thr Ser Thr Ile Thr Gly Asn Ser Tyr Thr Ala Asp Ser Asn
                410                 415                 420

ggc cag atc tgc gtg cag ctg tcc agc ggc gga ccc gag gcg ctc gtc      1456
Gly Gln Ile Cys Val Gln Leu Ser Ser Gly Gly Pro Glu Ala Leu Val
            425                 430                 435

gtc tcc acc gcg acc ggc acc gcc acc gcg acg act ctg tcc acg acc      1504
Val Ser Thr Ala Thr Gly Thr Ala Thr Ala Thr Thr Leu Ser Thr Thr
        440                 445                 450

acc aag acg tcc acc tcg acc gcc tcc tgc gcc gcc acc gtc gcc gtc      1552
Thr Lys Thr Ser Thr Ser Thr Ala Ser Cys Ala Ala Thr Val Ala Val
    455                 460                 465

acc ttc aac gag ctc gtc acc acg aac tac ggc gac acc atc cgc ctg      1600
Thr Phe Asn Glu Leu Val Thr Thr Asn Tyr Gly Asp Thr Ile Arg Leu
470                 475                 480                 485

acg ggc tcc atc tcc cag ctc agc agc tgg agc gca acc tcc ggg ctg      1648
Thr Gly Ser Ile Ser Gln Leu Ser Ser Trp Ser Ala Thr Ser Gly Leu
                490                 495                 500

gcc ctg agc gcg tcc gcg tac acg tcc agc aac ccg ctc tgg agc gtg      1696
Ala Leu Ser Ala Ser Ala Tyr Thr Ser Ser Asn Pro Leu Trp Ser Val
            505                 510                 515

acg gtc agc ctg ccg gcc ggc acg tcg ttc gag tac aag ttc gtc cgc      1744
Thr Val Ser Leu Pro Ala Gly Thr Ser Phe Glu Tyr Lys Phe Val Arg
        520                 525                 530

atc acg agc gac ggc acc gtg acc tgg gaa tcg gac ccg aac cgc agc      1792
Ile Thr Ser Asp Gly Thr Val Thr Trp Glu Ser Asp Pro Asn Arg Ser
    535                 540                 545

tac acc gtc ccg acg tgc gcg agc acc gcg acg atc agc aat acc tgg      1840
Tyr Thr Val Pro Thr Cys Ala Ser Thr Ala Thr Ile Ser Asn Thr Trp
550                 555                 560                 565

cgg tgagctctgg acgtgttgta catataggag gccgttgaga ggccggggcg           1893
Arg

gttggtggtc ggggtgaatg gggggttgat gctttttcgt tgtgtcggtg aga           1946
```

```
<210>  2
<211>  586
<212>  PRT
<213>  Chaetomium sp.

<400>  2

Met Arg Ser Phe Leu Ala Leu Ser Ala Leu Leu Leu Leu Tyr Pro Leu
-20                 -15                 -10                 -5

Gln Leu Leu Ala Ala Ser Asn Ser Asp Trp Arg Ser Arg Asn Ile Tyr
            -1  1               5                   10

Phe Ala Leu Thr Asp Arg Val Ala Asn Pro Ser Thr Thr Thr Ala Cys
```

15 20 25

Ser Asp Leu Ser Asn Tyr Cys Gly Gly Thr Trp Ser Gly Leu Ser Ser
30 35 40

Lys Leu Asp Tyr Ile Gln Gly Met Gly Phe Asp Ser Ile Trp Ile Thr
45 50 55 60

Pro Val Val Glu Asn Cys Asp Gly Gly Tyr His Gly Tyr Trp Ala Lys
65 70 75

Ala Leu Tyr Asn Val Asn Thr Asn Tyr Gly Ser Ala Asp Asp Leu Lys
80 85 90

Asn Phe Val Ala Ala Ala His Ala Lys Gly Met Tyr Val Met Val Asp
95 100 105

Val Val Ala Asn His Met Gly Ser Cys Gly Ile Ala Asn Leu Ser Pro
110 115 120

Pro Pro Leu Asn Glu Gln Ser Ser Tyr His Thr Gln Cys Asp Ile Asp
125 130 135 140

Tyr Ser Ser Gln Ser Ser Ile Glu Thr Cys Trp Ile Ser Gly Leu Pro
145 150 155

Asp Leu Asp Thr Thr Asp Ser Thr Ile Arg Ser Leu Phe Gln Thr Trp
160 165 170

Val His Gly Leu Val Ser Asn Tyr Ser Phe Asp Gly Leu Arg Val Asp
175 180 185

Thr Val Lys His Val Glu Lys Asp Tyr Trp Pro Gly Phe Val Ser Ala
190 195 200

Ala Gly Thr Tyr Ala Ile Gly Glu Val Phe Ser Gly Asp Thr Ser Tyr
205 210 215 220

Val Ala Gly Tyr Gln Ser Val Met Pro Gly Leu Leu Asn Tyr Pro Ile
225 230 235

Tyr Tyr Pro Leu Ile Arg Val Phe Ala Gln Gly Ala Ser Phe Thr Asp
240 245 250

Leu Val Asn Asn His Asp Thr Val Gly Ser Thr Phe Ser Asp Pro Thr
255 260 265

Leu Leu Gly Asn Phe Ile Asp Asn His Asp Asn Pro Arg Phe Leu Ser
270 275 280

Tyr Thr Ser Asp His Ala Leu Leu Lys Asn Ala Leu Ala Tyr Val Ile

285 290 295 300

Leu Ala Arg Gly Ile Pro Ile Val Tyr Tyr Gly Thr Glu Gln Gly Tyr
305 310 315

Ser Gly Ser Ser Asp Pro Ala Asn Arg Glu Asp Leu Trp Arg Ser Gly
320 325 330

Tyr Ser Thr Thr Gly Asp Ile Tyr Thr Thr Ile Ala Ala Leu Ser Ala
335 340 345

Ala Arg Thr Ala Ala Gly Gly Leu Ala Gly Asn Asp His Val His Leu
350 355 360

Tyr Thr Thr Asp Asn Ala Tyr Ala Trp Ser Arg Ala Ser Gly Lys Leu
365 370 375 380

Ile Val Val Thr Ser Asn Arg Gly Ser Ser Asp Ser Ser Thr Ile Cys
385 390 395

Phe Ser Thr Gln Gln Ala Ser Gly Thr Thr Trp Thr Ser Thr Ile Thr
400 405 410

Gly Asn Ser Tyr Thr Ala Asp Ser Asn Gly Gln Ile Cys Val Gln Leu
415 420 425

Ser Ser Gly Gly Pro Glu Ala Leu Val Val Ser Thr Ala Thr Gly Thr
430 435 440

Ala Thr Ala Thr Thr Leu Ser Thr Thr Thr Lys Thr Ser Thr Ser Thr
445 450 455 460

Ala Ser Cys Ala Ala Thr Val Ala Val Thr Phe Asn Glu Leu Val Thr
465 470 475

Thr Asn Tyr Gly Asp Thr Ile Arg Leu Thr Gly Ser Ile Ser Gln Leu
480 485 490

Ser Ser Trp Ser Ala Thr Ser Gly Leu Ala Leu Ser Ala Ser Ala Tyr
495 500 505

Thr Ser Ser Asn Pro Leu Trp Ser Val Thr Val Ser Leu Pro Ala Gly
510 515 520

Thr Ser Phe Glu Tyr Lys Phe Val Arg Ile Thr Ser Asp Gly Thr Val
525 530 535 540

Thr Trp Glu Ser Asp Pro Asn Arg Ser Tyr Thr Val Pro Thr Cys Ala
545 550 555

Ser Thr Ala Thr Ile Ser Asn Thr Trp Arg

560 565

<210> 3
<211> 18
<212> DNA
<213> Artificial

<220>
<223> Primer amyD1

<220>
<221> misc_feature
<222> (3)..(3)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (12)..(12)
<223> n is a, c, g, or t

<400> 3
gsntaccayg gntactgg 18

<210> 4
<211> 17
<212> DNA
<213> Artificial

<220>
<223> Primer amyD2R

<220>
<221> misc_feature
<222> (15)..(15)
<223> n is a, c, g, or t

<400> 4
tarayratkg gratncc 17

<210> 5
<211> 17
<212> DNA
<213> Artificial

<220>
<223> Primer AM835n-s1

<220>
<221> misc_feature
<222> (15)..(15)
<223> n is a, c, g, or t

<400> 5
aayaartayt tygcnyt 17

<210> 6
<211> 17
<212> DNA
<213> Artificial

<220>

<223> Primer AM835n-s2a

<220>
<221> misc_feature
<222> (3)..(3)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (6)..(6)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (15)..(15)
<223> n is a, c, g, or t

<400> 6
ctnggngaya grgtngc 17

<210> 7
<211> 17
<212> DNA
<213> Artificial

<220>
<223> Primer AM835n-s2b

<220>
<221> misc_feature
<222> (3)..(3)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (6)..(6)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (15)..(15)
<223> n is a, c, g, or t

<400> 7
ctnggngayc grgtngc 17

<210> 8
<211> 17
<212> DNA
<213> Artificial

<220>
<223> Primer AM835n-s2c

<220>
<221> misc_feature
<222> (3)..(3)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (6)..(6)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (15)..(15)
<223> n is a, c, g, or t

<400> 8
ctnggngayc gygtngc 17

<210> 9
<211> 18
<212> DNA
<213> Artificial

<220>
<223> amy835as1

<400> 9
gcggatagta gatgggat 18

<210> 10
<211> 17
<212> DNA
<213> Artificial

<220>
<223> Primer amy835f2

<400> 10
gtgcgtcctt caccgat 17

<210> 11
<211> 17
<212> DNA
<213> Artificial

<220>
<223> amy835s00

<400> 11
tcccgtcatc ctctctt 17

<210> 12
<211> 18
<212> DNA
<213> Artificial

<220>
<223> amy835as01

<400> 12
tctcaccgac acaacgaa 18

<210> 13
<211> 20
<212> PRT
<213> Chaetomium sp.

<220>
<221> misc_feature
<222> (9)..(9)
<223> Xaa can be any naturally occurring amino acid

<400> 13

Ala Ser Asn Ser Asp Trp Arg Ser Xaa Asn Lys Tyr Phe Ala Leu Gly
1 5 10 15

Asp Arg Val Ala
20

<210> 14
<211> 17
<212> DNA
<213> Artificial

<220>
<223> amy835as3

<400> 14
gtagtcaatg tcgcact 17

<210> 15
<211> 18
<212> DNA
<213> Artificial

<220>
<223> amy835f1

<400> 15
ccatctacta tccgctca 18

<210> 16
<211> 28
<212> DNA
<213> Artificial

<220>
<223> Primer AM835.1

<400> 16
ccaggatccg tcaggatgcg atccttcc 28

<210> 17
<211> 29
<212> DNA
<213> Artificial

<220>
<223> Primer AM835.2

<400> 17
cgtctcgagg acacaacgaa aaagcatca 29